Award Number: W81XWH-12-1-0333

TITLE: Determining the Location of DNA Modification and Mutation Caused by UVB Light in Skin Cancer

PRINCIPAL INVESTIGATOR: Monica Ransom, PhD

CONTRACTING ORGANIZATION: University of Colorado Denver
Aurora, CO 80045

REPORT DATE: September 2015

TYPE OF REPORT: Final Annual Summary

PREPARED FOR: U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

1. REPORT DATE	2. REPORT TYPE	3. DATES COVERED
September 2015	Final Annual Summary	15 Aug 2012 – 14 Aug 2015

4. TITLE AND SUBTITLE

Determining the Location of DNA Modification and Mutation Caused by UVB Light in Skin Cancer

5a. CONTRACT NUMBER
W81XWH-12-1-0333

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
Monica Ransom, PhD

E-Mail: Monica.Ransom@ucdenver.edu

5d. PROJECT NUMBER

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

University of Colorado, Denver
Aurora, CO 80045

8. PERFORMING ORGANIZATION REPORT NUMBER

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)
U.S. Army Medical Research and Materiel Command
Fort Detrick, Maryland 21702-5012

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)

12. DISTRIBUTION / AVAILABILITY STATEMENT
Approved for Public Release; Distribution Unlimited

13. SUPPLEMENTARY NOTES

14. ABSTRACT

Extrinsic modifications to the DNA bases such as pyrimidine dimers can arise from a variety of exposures and can lead to aberrant cell growth or death. A detailed view of this base modification is necessary for a more complete view of genetic and epigenetic regulation but the process is poorly understood. We have developed a method to look at the precise genomic position of these modifications using a next generation sequencing approach. During this research period we have optimized this protocol to obtain libraries from UVB damaged human cells. We have purified or obtained the enzymes needed for aim 1 of this research. Using these enzymes we have shown that we obtain cleavage patterns consistent with the administered UV dosage and that sequencing libraries generated for both yeast and human cells show pyrimidine bias on the 5' end, indicating that we are sequencing the dimers. Understanding where these modifications occur is a critical first step to understanding the mutations they cause.

15. SUBJECT TERMS
UV DNA modification, HeLa cells, Skin Cancer

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(include area code)*
U	U	U	UU	19	USAMRMC

Table of Contents

INTRODUCTION

The exposure of skin cells to ultraviolet (UV) radiation from the sun damages DNA and leads to the formation of pyrimidine dimers. There are 2 main forms of pyrimidine dimers cyclobutane pyrimidine dimers (CPDs) and 6-4 photoproducts(6-4pps), which are bulky DNA adducts that prevent replication and transcription from occurring until they are repaired. Mutations that have hallmark profiles of UV damage have been found in a variety of genes such as p53 have been found in many skin cancers. While the study of these mutations than lead to skin cancer has been underway for many years there is much less known about the modification events that underlie these mutations. In this study we undertook the task of determining the genome-wide distribution of UV-induced DNA modifications, and to elucidate which of these modifications lead to eventual mutations via high-throughput sequencing approaches. Our lab developed a method of identifying DNA base modifications by combining commercially available base excision enzyme cleavage with next-generation sequencing. We have shown in these libraries that the sequences that are derived from pyrimidine dimer modification come from sequences that contain pyrimidine dimers as well as seeing the proportion of dimers is similar to that seen in other mapping strategies. These experiments show that mapping of UV dimer modification may yield insight into how and where these modifications are formed in DNA.

KEYWORDS: Pyrimidine dimers, UV light, modification mapping, excision repair, UVDE, cyclobutane pyrimidine dimers, 6-4 photoproducts

OVERALL PROJECT SUMMARY:

The primary task in the statement of work dealt with the generation of sequencing libraries in yeast and human cells. In previous work we have shown that libraries can be generated from yeast irradiated with UVC light using a commercial glycosylase and photolyases from a collaborating lab. During the process of transitioning to UVB light and human cell experiments the commercial glycosylase went off the market so the first step in this task became to generate these enzymes within the lab. UVDE is the S. pombe glycosylase that can digest both CPD and 6-4 dimers. We obtained a plasmid construct for S. pombe UVDE and purified protein over a glutathione column (Fig. 1A) (1). It was determined that the homemade enzyme worked equivalently to the commercial one when used as the same concentration (Fig. 1B). We next wanted to validate this enzyme for library preparation but were unable to obtain another sample of the photolyases we used for our preliminary results. We obtained constructs to make our own enzymes and went through several rounds of purification using amylose columns followed by both an S column as well as a heparin column. Although we were able to obtain relatively pure protein we were unable to validate enzyme activity through library preparation. After several tries at this we contacted a lab that purifies these enzymes for crystallization and were able to obtain a sample of both photolyases (*2, 3*).

Figure 1.

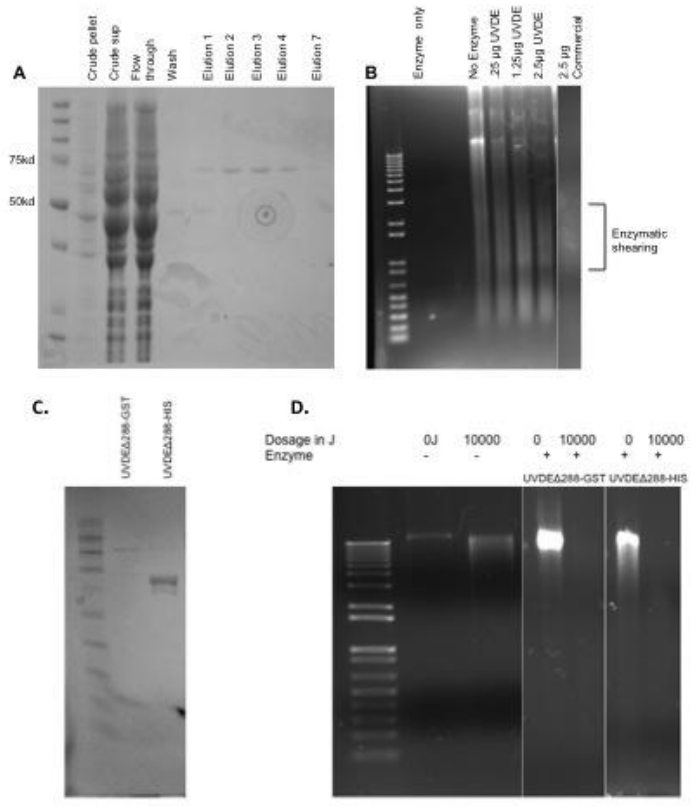

Figure 1. UVDE protein made in the lab works similarly to commercial UVDE enzyme from Trevigen. Protein was purified from yeast containing cup-1 promoter driven UVDEΔ288-GST after induction. Following purification the UVDE protein eluted in 10mM glutathione in fractions 2-4 as seen in Fig 1C. Homemade enzyme was compared to commercial enzyme for cleavage

of yeast DNA treated with 10000J/m2 of UV irradiation. When used at the same concentration (compare lanes 4 &5) we obtained similar shearing patterns as outlined to the right (Fig 1D).

We used our homemade UVDE as well our newly obtained photolyases and made sequencing libraries with highly irradiated yeast DNA to confirm our preliminary findings in a biological replicate. DNA was sheared using our UVDE enzyme (Fig. 2A) and after treatment with either the CPD photolyase or the 6-4 photolyase, to repair the ends, sequencing libraries were obtained and run on the Illumina platform (Fig. 2B). We compared our data to a sheared control as well as the whole genome dinucleotide pattern and determined that there was a bias at the 5' end of our sample for dipyrimidines as expected if we are generating a cleavage event at damaged bases. This bias was similar if not as robust as that seen previously with the old enzymes **(**compare Fig. 2c to Fig. 2d). We went on the further improve the UVDE protein preparation by replacing yeast with E. coli expression. We used gateway to clone the S. pombe UVDE glycosylase with the delta 288 mutation (1) into a pet-53-His vector under the T7 promoter. We transformed this construct into E. coli that are competent for protein expression and induced them overnight in .4mM IPTG. The cells were harvested, frozen and lysed by sonication. The lysate was clarified by centrifugation and the supernatant was purified over a nickel column as compared to the initial yeast protein purification (Fig. 1C). This protein was concentrated and compared to our previous yeast purification and found to yield 10-15 times as much protein as the previous technique. The enzyme still sheared efficiently as shown in (Fig. 1D).

Figure 2.

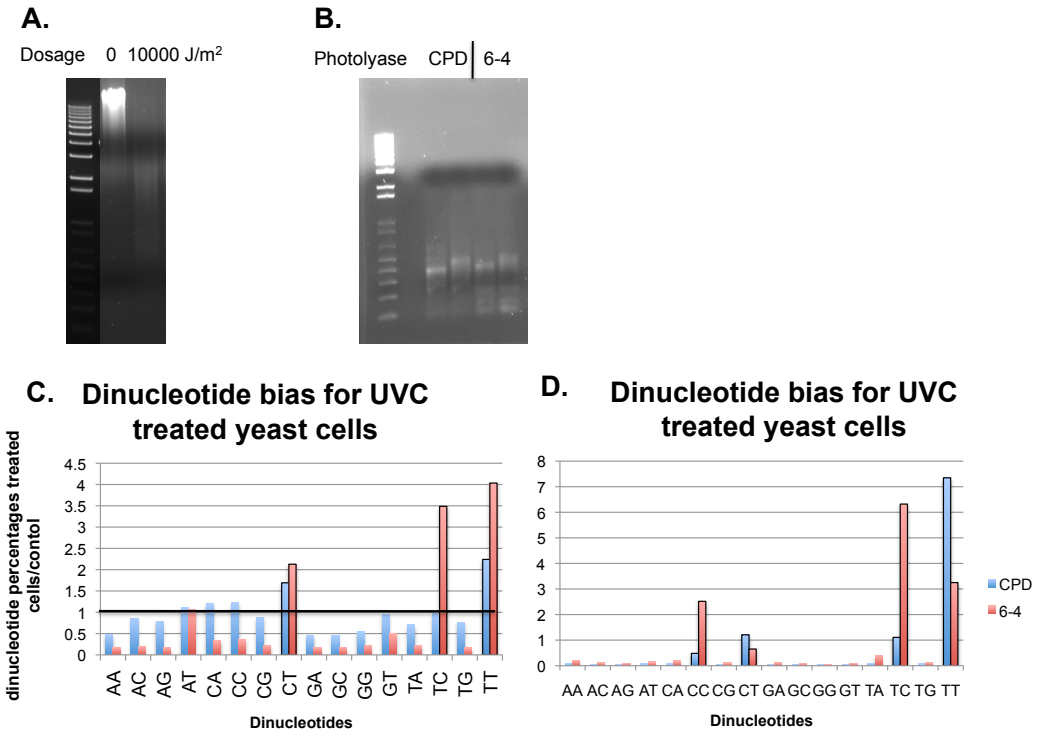

Figure 2. Illumina sequencing libraries were obtained from yeast cells treated with a high dose UVC light. Yeast cells were treated with UVC light at 0 and 10000 J/m2 and genomic DNA preps were analyzed for cleavage with UVDE (Fig. 2A). Samples were treated with CPD or 6-4 photolyase and then run through standard Illumina prep and the libraries were obtained for 2 size-selected fractions (Fig 2B). Dinucleotide bias on the 5' end of sample reads as compared to

control dinucleotide bias is shown (Fig 2C). The biased dipyrimidines are outlined in black for comparison. Dinucleotide bias from a sample prepared with commercial UVDE and photolyases from Aziz Sancar (4) is shown for comparison (Fig. 2D).

With a validated glycosylase in hand we transitioned into dosing human HeLa cells. We decided to initially look at UVC to troubleshoot any problems that may occur with samples that contained more damage. We began by looking at the lethality associated with UV irradiation. We looked at both yeast and human cells. Yeast cells were irradiated at a given dose in a Statalinker and confirmed using a UVP dosimeter. Cells were plated onto rich media and the irradiated sample counts were normalized to an unirradiated control (Fig 3A). We also looked in HeLa cells using lower doses because it has been shown that human cells cannot tolerate high doses of UV irradiation (*5,6*). We irradiated cells in PBS and seeded fresh plates for 24 hours in DMEM before scoring with trypan blue exclusion for viability. Counts were normalized to an unirradiated control to account for normal cell death (Fig. 3B). We saw UV50 lethality in yeast at approxiametly 500J/m^2 and in HeLa cells the UV50 dose was 60J/m^2 .

Figure 3.

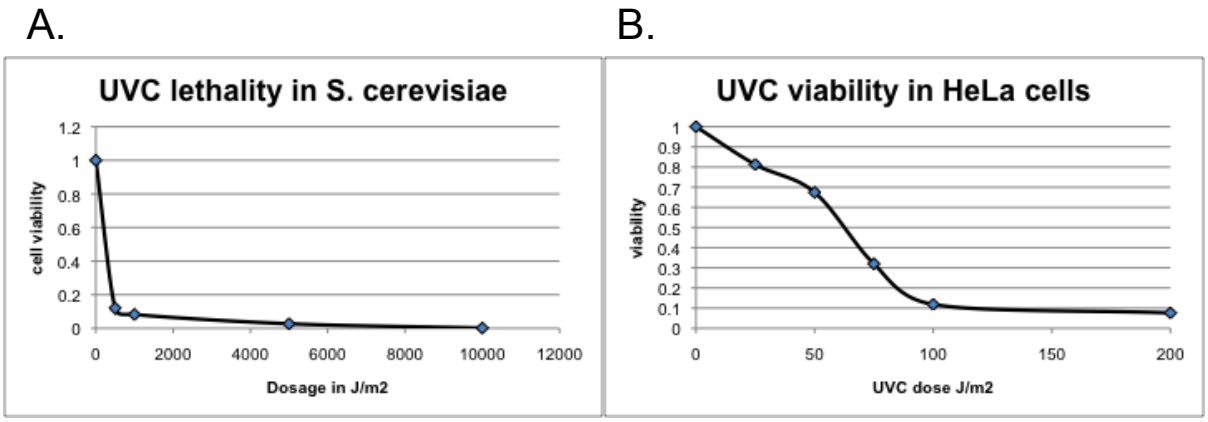

Figure 3. Human cells are 10 times more sensitive to UVC than yeast cells. Yeast cells were irradiated for the given doses and plated onto rich media at a known density. Cells were allowed to outgrow for 2 days and scored for colony formation and normalized to an unirradiated control as shown in Fig 3A. Human cells were irradiated for the given doses and plated in 6 well plates to recover for 24 hours. Cells were then trypsinized and scored for viability using trypan blue exclusion and normalized to unirradiated cells as shown in Fig. 3B.

We further wanted to determine the cleavage pattern of UVDE in irradiated human cells. We grew HeLa cells to confluency and irradiated at a given dose with UVC light as measured by a spectrophotometer. Genomic DNA was isolated using a gentle protocol and 2 µg of DNA for each dosage was cleaved with 1.5 µg of UVDE for 4 hours at 30˚ and ran on a gel (FIG. 4A). Cleavage to lower molecular weight fragments was seen starting at 1000J/m2 and DNA degradation was beginning to occur at 20000 J/m2. We next wanted to optimize photolyase cleavage and library preparation. We took three of our samples 0 J/m2, 500 J/m2 (low dose) and 10000J/m2 and digested with UVDE (Fig 4B). We then treated these samples with either CPD photolyase of 6-4 photolyase for 1 hour under UVA light. Samples were then run through standard Illumina preparation including polishing, a-tailing, adapter ligation and PCR. Libraries were obtained for both the low dose and the high dose samples but the low dose samples were in low abundance (Fig. 4C).

Figure 4.

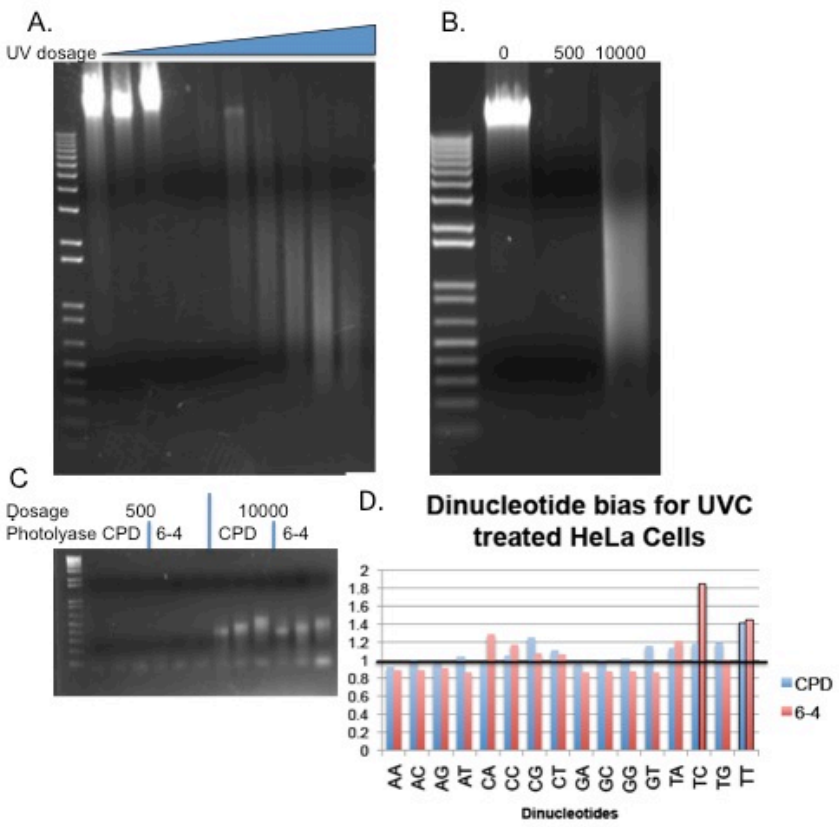

Figure 4. Illumina sequencing libraries were obtained from HeLa cells treated with low and high dose UVC light. HeLa cells were treated with UVC light in increasing doses from 0 to 20000 J/m2 and genomic DNA preps were analyzed for cleavage with UVDE (Fig. 4A). Samples from 3 dosages 0, 500, and 10000 were scaled up (Fig. 4B) and treated with CPD or 6-4 photolyase. Cells were then run through standard Illumina prep and the libraries were obtained for several size-selected fractions (Fig 4C). Samples from HeLa cells were treated with 10000 J/m2 of UVC light, sheared with UVDE, repaired with either CPD or 6-4 photolyases, and made into Illumina libraries. Dinucleotide bias on the 5' end of sample reads as compared to control dinucleotide bias is shown (Fig. 4D). The biased dipyrimidines are outlined in black for comparison.

Libraries were pooled and sequenced on the Illumina MiSeq platform. We obtained approximately 12.5 million combined reads. For the low dose libraries a low percentage of the reads aligned to the hg18 build of the human genome. This is generally a sign of low library quality and is not surprising due to the weak shearing and PCR bands. These libraries also showed no bias for dinucleotides indicating that they are not adequate UV libraries (Data not shown). The high dose UV libraries aligned much better at 71% for CPD and 73% for 6-4 libraries that is typical for human libraries (7). These reads were then processed to determine the dinucleotide composition on the 5' end. The percentage of each dinucleotide combination for the whole human genome was then determined and used as a control for base bias in the genome. The data was plotted as the percentage of each dinucleotide ratio in UV irradiated DNA/ dinucleotide ratio of the control sample. Dinculeotide bias was found in the 1st base of the

read and the base previous to it as expected considering that one base of the dinucleotide was cleaved during the photolyase repair step (Fig 4D) (8). Data from irradiated yeast treated with the same enzymes and protocols for comparison (Fig 2C). Although the bias is significantly reduced compared to the yeast sample, it is present and we continued to streamline the approach to improve our method.

To begin to streamline our data we wanted to better understand the sensitivity of our method. We went back to the original data for yeast, which had the best dipyrimidine bias and did additional analysis. We determined that the sensitivity of this assay in yeast was quite high with more than 85% of the aligned sequences acquired deriving from genomic positions with pyrimidine dimers. In total we saw that 38% of the total genomic dipyrimidines were hit in the CPD library with 72% of the TT dipyrimidines in the genome having reads. The 6-4pp library hit only 5% of the total dipyrimidines indicating more specificity of the damage itself or of the repair enzyme used to generate the libraries. We also went on to look and the average number of hits in the two libraries and subsequently saw a increase in the average number of times each hit occurred in 6-4 photoproduct libraries, again indicating an increased specificity. We went on to further look at the local base content surrounding the modified dipyrimidines and saw that in CPD libraries the bases up and downstream of the modified base reflected the same percentages as the yeast genome (Fig. 5A), whereas in the 6-4pp library the base 3' to the dipyrimidine shows a bias to being an A residue (Fig 5B) (9). This may indicate an otherwise unknown specificity for the damage to occur within these trinucleotides or for the repair enzymes to be less efficient at repair of these sites. We also further analyzed the genomic positions of this data and showed that the coverage of modifications was generally uniform across the genome in yeast and the location of the dipyrimidines couldn't be associated with chromatin context or several other DNA features tested (data not shown).

Figure 5.

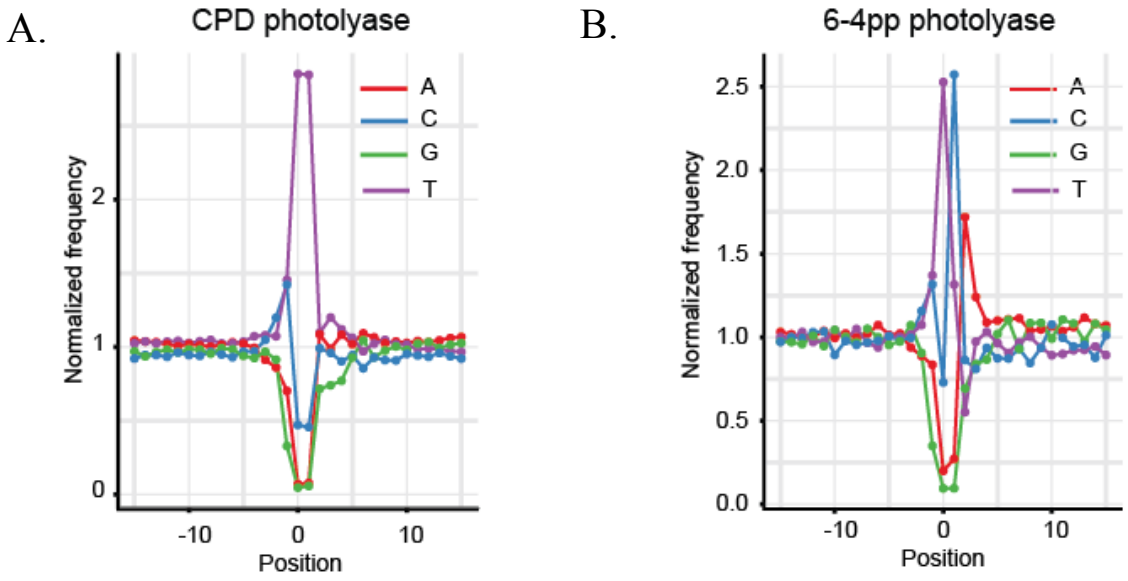

Figure 5. Additional data analysis on libraries obtained from yeast cells treated with a high dose of UVC light. Frequency of nucleotides relative to mapped positions of sequences from pre-digestion Excision-seq libraries for mapping cyclobutane dimers in S. cerevisiae. Position 0 corresponds to the mapped position of the 5' end for CPD (Fig. 5A) and 6-4 libraries (Fig. 5B).

Now that we have the appropriate enzymes and have libraries showing some bias the protocol needs to be streamlined to use UVB light at more biologically relevant dosages. To this end we obtained a UVB light from Coleman and began performing experiments but upon measuring the UV wavelength with a dosimeter determined that the UV spectrum was quite broad and all three wavelengths of UV light were being administered. To address this we obtained an LED bulb from Qphotonics that emits light at 315nm ± 10 nm (10) and incorporated it into a light source that emits UVB at 20J/m^2s. Using this light source with primary keratinocyte cells we were able to show low levels of DNA damage as measured by UVDE cleavage (Fig 8A). This mild shearing pattern is obtained because the DNA damage is not saturated enough to yield smaller molecular weight fragments.

Upon seeing the low amount of shearing in a biologically relevant UVB dosage, we decided to troubleshoot our protocol using low doses of UVC light in yeast cells. When the UV dosage is lowered we see a decrease in the percentage of 5' biased ends in our sample libraries below 5000J/m^2. This is due to the lack of sufficiently small double stranded DNA fragments that have dimers on either end. This also leads to an increasing level of background noise from other DNA breaks that are occurring in the cells or during the processing of the DNA. To work around this we developed a circular ligation approach that allows us to map single modifications as well as to remove the bias generated during the PCR step (Fig. 6).

Figure 6.

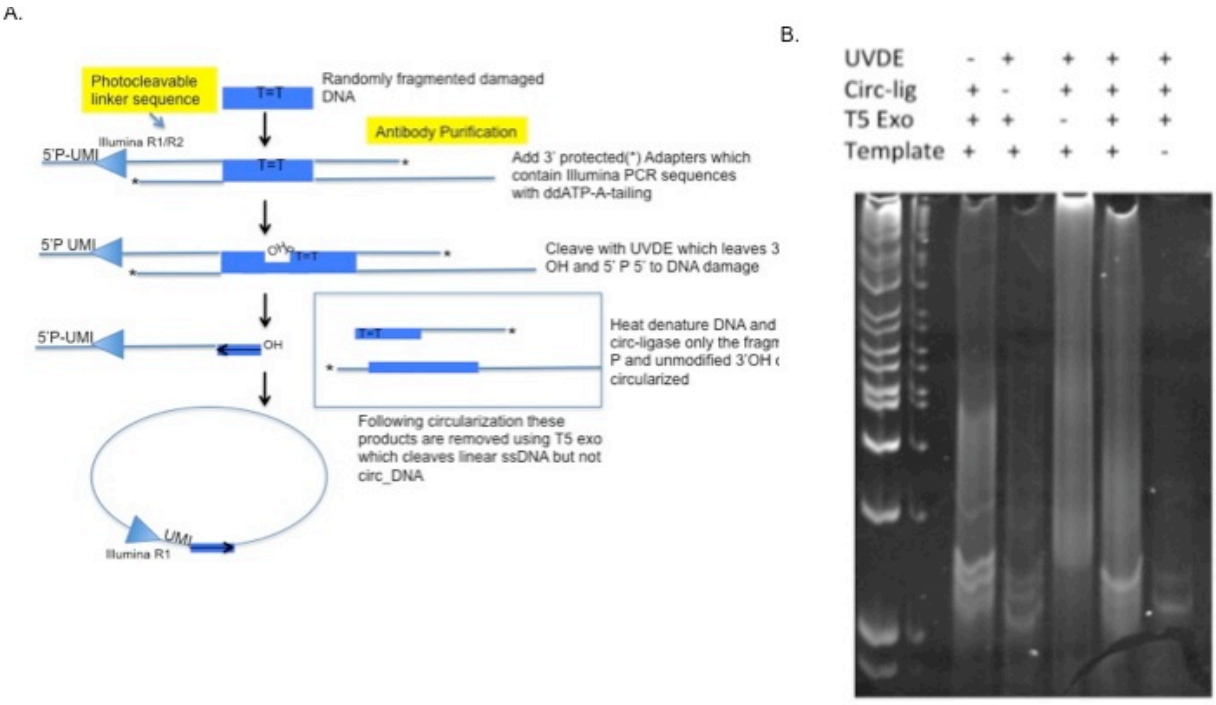

Figure 6. Scheme for circularization protocol. To generate pyrimidine dimer specific libraries sheared DNA containing photodimers is ligated to an Illumina circular adapter containing a UMI and cleaved with UVDE. This cleavage event leaves a 3' OH that can circularize in the presence of Circ-Ligase (Epicentre). Non-photodimer specific fragments are removed with T5 exonuclease (Invitrogen) prior to circularization. The circular fragment can then be PCR amplified with standard Illumina adapters. Additional protocol changes indicated in yellow are adding an antibody purification step after shearing to increase the pool of photodimer containing

sequences in our libraries, and adding a photocleavable linker into our adapter that can prevent circular PCR amplicons (Fig. 6A). Representative Circ-Ligase preparation with no UVDE, no Circ-Ligase, and no template negative controls as well as a no T5 positive control (Fig.6B). Lane 4 indicates a sequencing library that has a signal similar to that of the positive control lane 3.

Figure 7.

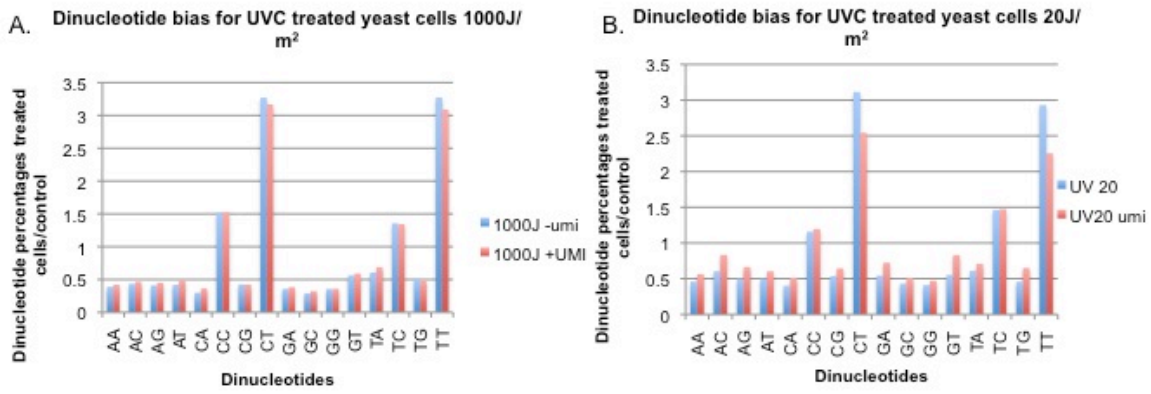

Figure 7. Pyrimidine dimers are enriched at the 5' ends in low dose UV damaged libraries. Yeast cells were irradiated with either 1000J/m^2 (Fig. 7A) or 20J/m^2 (Fig. 7b) of UVC light and DNA was isolated and prepared using the protocol described previously. In all samples we determined the percentage of the dinucleotides at the 5' of libraries between a UV damaged library and the dinucleotides present in genomic DNA. All 4 dinucleotides show enrichment in the UV treated sequencing library. The blue bars indicate the data prior to accounting for the UMI derived PCR bias the red following it.

Using this approach we generated libraries for UVC treated yeast cells at dosages of 1000J/m^2 and 20J/m^2 (Fig 7A and B). These libraries showed bias at a lower UV dosage indicating that achieving low dose UVB libraries from human cells would be possible. In Fig. 7 we show that the unique molecular identifier in these adapters can be used to remove PCR duplicates (11). This is done by introducing a 12 base pair random sequence into the adapter that is read at the beginning of the sequencing read. These sequences act as a barcode for each ligation event and any non-unique sequences indicate a PCR duplication and not a unique ligation of a DNA molecule. Using this technique we removed a small subset of our sequences that are PCR duplicates and reanalyzed the data. We obtained a similar trend that showed that the bias we are seeing is not due to PCR amplification and that most of the sequences came from unique pyrimidine dimers.

Figure 8.

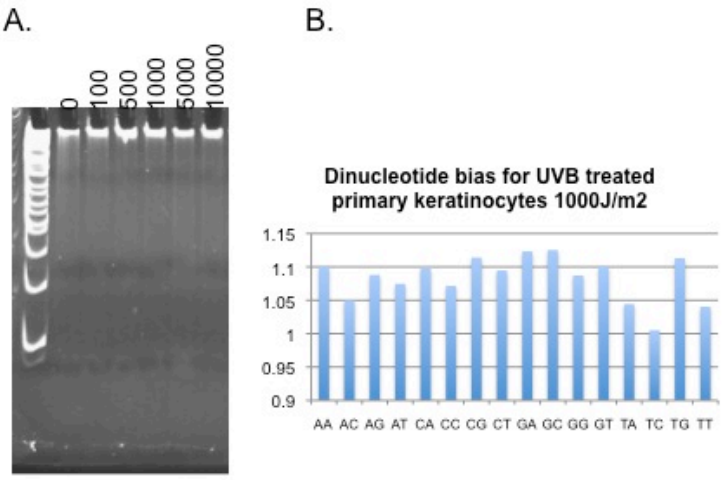

Figure 8. UVB light generates DNA damage that is visible following UVDE cleavage but is unable to form biased pyrimidine libraries. Primary keratinocytes were treated with the various dosages of damage indicated above the gel in $J/m^{2.}$ As the dosage increased the DNA fragmentation was increased in the smaller molecular weight ranges (Fig. 8A). This shearing is significantly less than seen with UVC dosages, as UVB is 100 fold less damaging (5). When this DNA is made into an Illumina library using the circularization protocol there is no dipyrimidine bias seen (Fig. 8B)

When UVB libraries were generated from human cells, we were unable to see clear DNA bias in several different dosages (Fig 8B). We believe this may be due to several causes such as background levels of single stranded breaks present in the DNA, mild shearing during the preparation of the DNA, or inefficient circular ligation. To address this further modifications were added to our protocol including adding an antibody purification step after shearing to increase the pool of photodimer containing sequences in our libraries, and adding a photocleavable linker into our adapter that can prevent circular PCR amplicons (Fig. 6a yellow).

When looking at the human libraries we saw a lot of high molecular weight bands indicating that we may be getting circle PCR going around multiple times (data not shown). To try to address this problem we added a photocleavable linker to the adapter sequence between the forward and reverse primer sequences to prevent the PCR from going multiple rounds (Fig. 6A). After the T5 exonuclease reaction we can cleave this sequence with UVA light to break the circle and to prevent circle amplification. After this modification we still did not obtain specific human cell dipyrimidine libraries (data not shown).

To try to troubleshoot the poor library specificity we looked at the various stages of the process using a control DNA to determine the steps of the process that are failing **(Fig. 9**).

Figure 9.

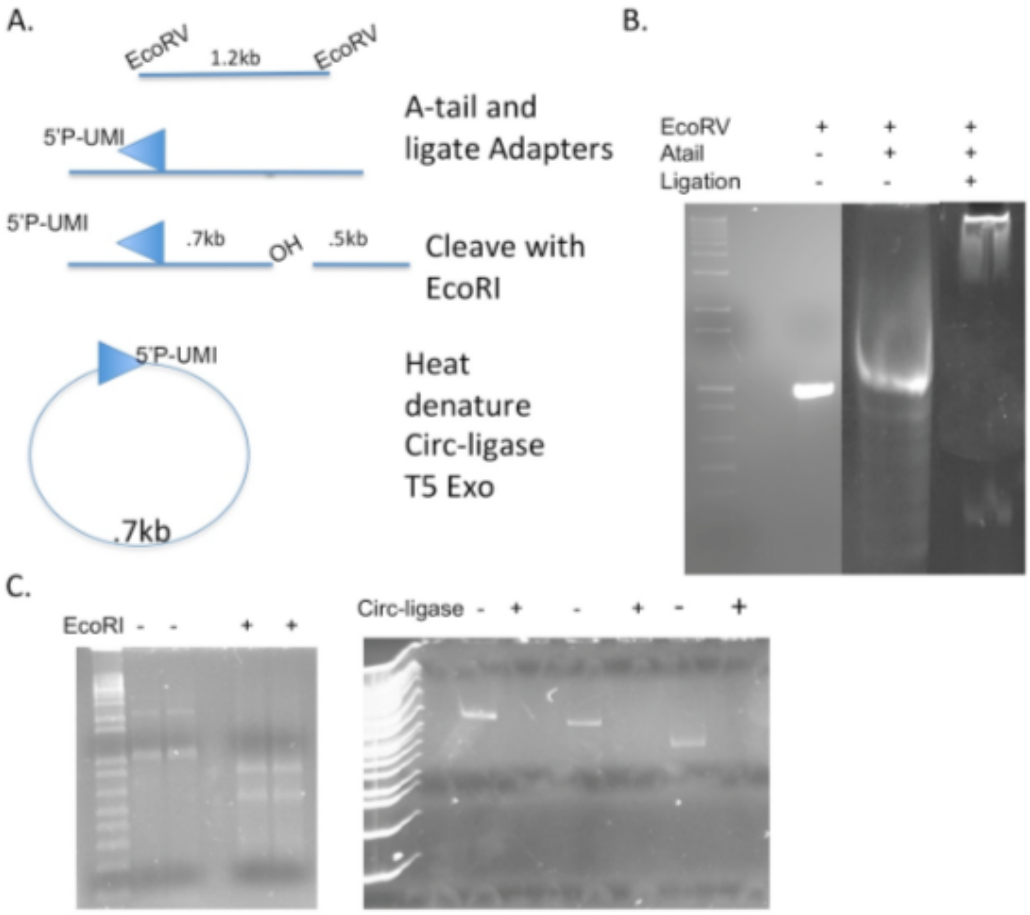

Figure 9. General scheme of the protocol we used to determine the problems in the library preparation. To troubleshoot the Circ-Ligase protocol we isolated a 1.2kb fragment from a vector using EcoRV. We then A-tailed and ligated the Illumina adapter (Fig. 9A). We can see than when we A-tail and ligate the adapter we generate a smeared high molecular weight band at high efficiency indicating that these two steps are working. We then digest with EcoRI to generate a free 3' OH and isolated the individual bands of which the 700 base pair fragment should be able to circularize. After treating with Circ-Ligase we see no circularization indicating that this is the step that is failing (Fig. 9b).

Using this approach we determined that it was the low efficiency of Circ-Ligase on these types of templates that was mostly to blame for poor library pools. To address this we used 2 methods; first to isolate damaged DNA and enrich our pool of available substrate for the ligase to work on, and second to purify an enzyme that may have Circ-Ligase activity for use at high concentrations to improve the enzyme efficiency.

To enrich the pool of available substrates for the Circ-Ligase reaction we took an immunoprecipitation approach using a commercially available antibody against CPD DNA damage. To test this antibody we used a dot blot to determine the amount of damage needed to detect with the antibody (Fig. 10A and B).

Figure 10.

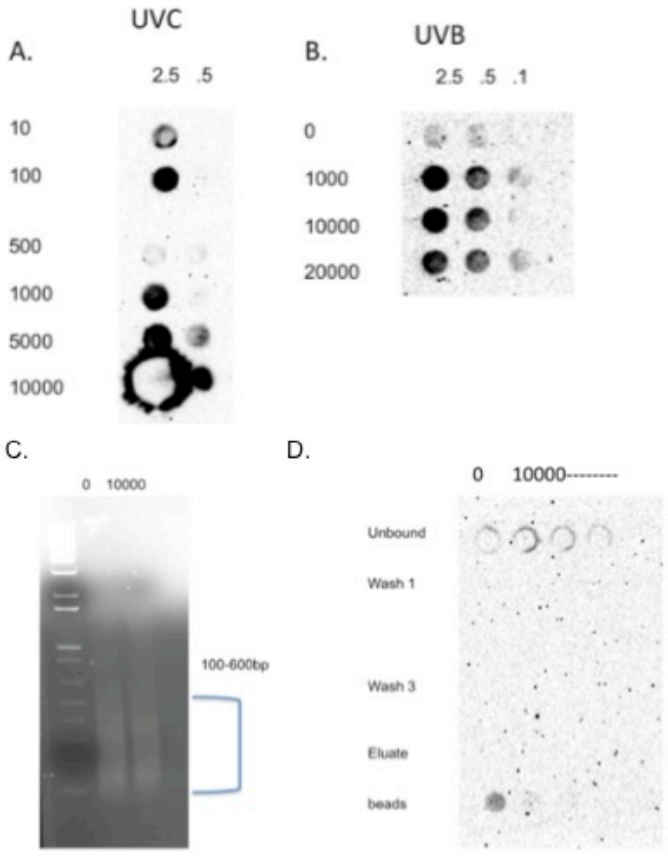

Figure 10. Anti-CPD antibody binds to UVC and UVB damaged DNA but fails to precipitate it. UVC and UVB damaged DNA (in J/m2) was denatured with NaOH and heat and applied to the membrane in the µg quantities indicated and fixed by crosslinking. The membrane was blocked in milk and probed with anti CPD primary antibody and anti-mouse HRP secondary antibody. We were able to see both UVB and UVC damage at .5µg even with low damage amounts (Fig. 10A and B). DNA from unirradiated and DNA irradiated at 10000J/m2 was sheared with a biorupter to yield fragments between 100-600 base pairs in (Fig. 10C). After binding DNA to anti-CPD antibody we immunoprecipitated with Dynabeads protein G. We took samples of the unbound DNA and from each wash as well as the eluate and what was left on the beads (Fig. 10D). We saw that there was signal in the unbound fraction and left on the beads. This is probably due to the non-specific binding of the antibody. It is unclear where the damaged DNA was lost during the precipitation.

We saw that with UVB and UVC we could detect DNA as low as .5µg at relatively low doses. We also saw low levels of detection in the 0J/m2 dosage, indicating that there may be some non-specific DNA binding with this antibody. We sheared 10µg of genomic DNA using a biorupter for 15 minutes on high with 30 seconds on and 30 seconds off. This yielded DNA that was 200bp-1kb in size (Fig. 10C). The sheared DNA was added to the antibody and incubated for 30 minutes at room temperature. The supernatant was removed as the unbound fraction, and the beads were washed with buffers of varying stringency, and the DNA was eluted with TE/1% SDS at 65˚ for 15 minutes. After several attempts at this protocol we were never able to

get the antibody-protein complexes to show up on the blot and the only signal we saw was most likely due to nonspecific antibody binding (Fig. 10D). After several attempts to get this protocol to work we moved on to the second method to improve library preparation.

To generate a homemade enzyme with Circ-Ligase activity we used a temperature sensitive allele of RNL-1 (12). We obtained the DNA sequence for this enzyme and using gateway cloning tagged it with 6x-his and engineered a stop codon. Using this construct we were able to obtain large quantities of the enzyme and show it had similar activity to the commercial enzyme (DNS). We are currently using this enzyme in much higher quantities to try to generate higher quality sequencing libraries.

Since we have as yet been unable to generate additional libraries from human cells damaged with UVB I went back to the one sample that showed bias in HeLa cells and probed the data further to try to determine if there was any additional information we could learn using the data we already had. To try to determine if there was any pattern of where the photodimers are in human cells we took the one library that showed bias from HeLa cells and used the data to intersect with segmentation annotations from human hepatocytes (13).

Figure 11.

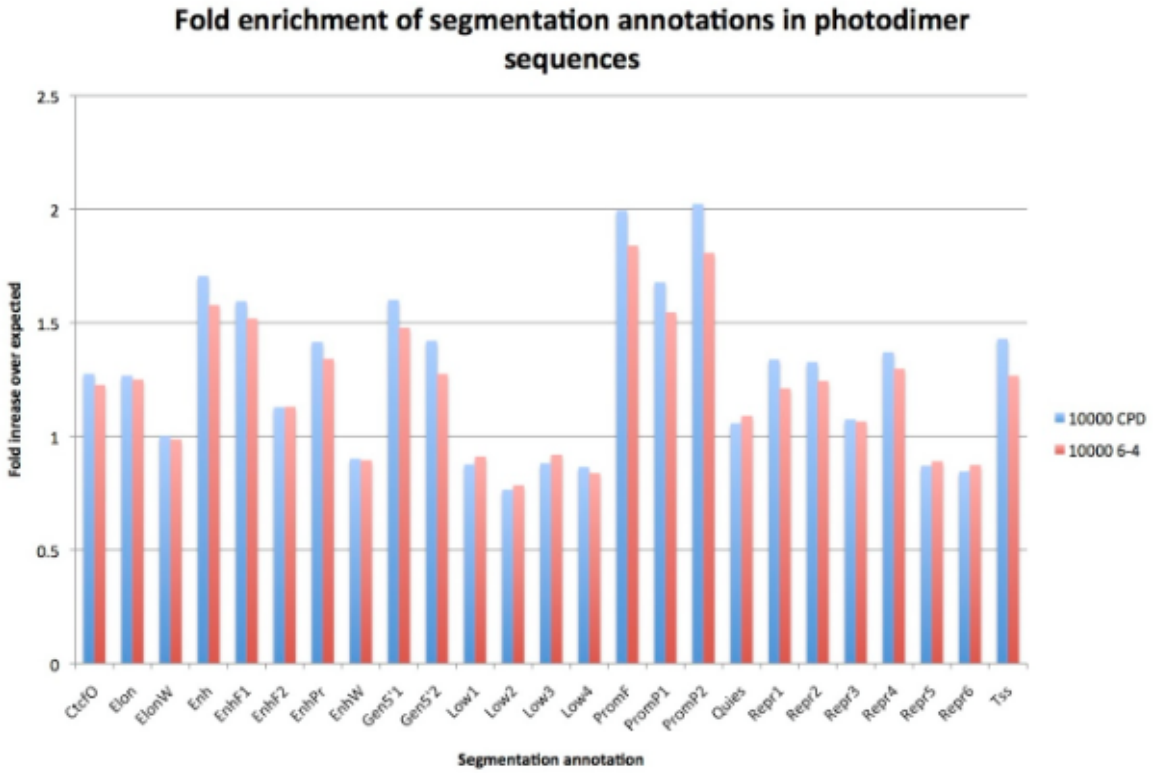

Figure 11. UVC induced photodimers are slightly enriched in promoter and enhancer regions. Data from HeLa cells treated with 10000 J/m2 of UVC light were aligned to the human genome. The mapped reads were converted to a bed file and intersected with human hepatocyte segmentation data (5). Fold enrichment of segmentation annotation as compared to expected is show. Enhancer and promoter regions are slightly enriched.

This type of analysis allows us to see which segmentation annotations are enriched in our library over the expected. Using this analysis we see slight enhancement in enhancer (Enh) and promoter (Prom) regions (Fig. 11). This is an expected pattern as these are the regions of the genome with open chromatin and it is expected that DNA damage is going to occur more

frequently in these regions. The enhancement seen is low most likely due to the fact that the library bias is not very strong so the background level is probably high. Once higher quality libraries are obtained this and similar analyses can be performed to understand where these modifications are occurring and which ones go on to cause mutations.

Tasks 3 and 4 related to generating SMRT libraries and sequencing them were not undertaken in this project due to a variety of technical difficulties including our department not obtaining and SMRT sequencer. The generation of the novel adapters necessary for this type of sequencing as well and the in depth troubleshooting required when dealing with DNA base modification library preparation further discouraged this analysis. It was determined that we would try to focus on the first tasks using protocols we are more familiar with rather than trying to generate samples for a new and unfamiliar system.

Task 5 was not performed because experiments in task 1 and 2 indicated high levels of UV damage were not being obtained and that there was a relatively high background in the samples we did obtain. This would have made the already difficult task of mutation calling impossible even with relatively high coverage. Also we spent some time trying to develop adapters specific to accurately identifying mutations from a small sample size and were unable to generate libraries with them. Taken all together we decided to not perform this task due to its high cost and poor potential outcome.

KEY RESEARCH ACCOMPLISHMENTS:

- Yeast UVC libraries contain photodimers at the expected ratio and are highly specific.
- 6-4 photoproducts contain a unique bias for an A in the 3' position.
- Circular ligation allows analysis of samples with low levels of DNA damage.
- Human cells treated with UVC contain low levels of DNA damage and are not captured well by the excision-seq method.

REPORTABLE OUTCOMES:

Manuscripts, abstracts presentations

D. Bryan*, M. Ransom*, B. Adane, K. York, J. Hesselberth, High resolution mapping of modified DNA nucleobases using excision repair enzymes, Genome Res. 24:1534.

Ransom, M and Hesselberth, J. Poster session presented at: Cell Biology and Molecular Oncology Retreat; 2013. Oct. 21; Aurora CO.

Ransom, M and Hesselberth, J. Poster session presented at: 5th annual postdoctoral research day; 2014 Mar. 14; Aurora, CO.

Patents and licences Nothing to report

Degrees obtained None

- Coursera courses
  - Learn to Program; The fundamentals
  - The Data Scientist's Toolbox
  - Programming for Everyone (Python)
  - R programming
  - Getting and Cleaning Data
  - Introductory Human Physiology
  - Exploratory Data analysis

  - UC Denver courses
  - MOLB 7621 Genome Analysis Workshop

Development of cell lines, tissue, or serum repositories

GSE51361-NCBI Gene Expression Omnibus

Funding applied for None

Employment and research opportunities applied for based on training None

CONCLUSION:

To begin performing the tasks outlined in the statement of work we first had to generate and test all of the new enzymes we have made and borrowed to determine if they work as well as the commercial enzymes that went off the market. To this end we made and sequenced a library under similar conditions to that we have made before and saw that while the library was not as robust the general trends and patterns remained the same.

During the tenure of this project we have spent significant time trying to modify our original UV mapping protocol to generate more specific libraries and libraries that contain low levels of DNA damage. To accomplish this task we began by generating a circular ligation approach that allows us to map single base modifications instead of relying on multiple modifications in a small region to generate libraries. This protocol allowed us to map yeast libraries that had low UV dosage. Using this protocol we tried to generate libraries in human cells. In several initial libraries we saw no bias of dipyrimidine ends. We tried several additional techniques to try and generate specific human libraries.

Initially we tried to modify the circ-ligase approach by adding a photocleavable linker in between the primer binding sites to prevent the high molecular weight species present in our libraries that we attributed to multi-circle PCR products. Libraries made using these new adapters yielded similar libraries to those we had seen previously. We also tried improving the libraries by enriching the dipyrimidine containing DNA that was present for ligation. To do this we obtained an anti-CPD antibody and showed that it bound to damaged DNA, but in several attempts to immunoprecipitate the damaged DNA we were never able to get efficient pull-down to try to make libraries.

We next decided to see at what stage the circ-ligase protocol was failing. We generated an artificial system that allowed us to look at each step in the protocol. We were able to determine that the A-tailing and adapter ligation reactions were efficient and seem to be going to completion. As expected we determined it was the circular ligation reaction that was inefficient with virtually no ligated product visible following ligation. This is a problem we were expecting since circ-ligase is designed to work in a very small volume with limited template and we are trying to use it in a larger volume with a lot of template. To try to address this we are trying to purify a temperature sensitive allele of RNL-1 that has been shown to have circular ligase activity. By purifying this enzyme we can obtain large quantities of enzyme that can be used at high concentration to generate libraries. We have generated this enzyme and it circularizes a control template. In the future we will use this enzyme to try again to obtain specific libraries in human cells at low doses.

We went back to some of the initial data we obtained from HeLa cells that had been dosed with UVC. We looked at this data again to try to determine if there was any pattern of where these modifications were occurring and determined that there was modest enhancement in promoters and enhancer regions. This is expected since these regions of chromatin are more open and likely to obtain damage. In the future we can look at this data in comparison to other datasets to try to determine other patterns of damage localization. We hope that this protocol can continue to be improved to allow us to further understand where these modifications are taking place.

We have generated a method to study the DNA modifications caused by exposure to UV light. We have shown that these libraries from high doses are more prevalent in regions of open chromatin. These new methods for studying genome wide distribution of UV modification may bring clarity to the relationship between UV DNA modification and mutation. We hope that with this new knowledge will come advancements in the prevention and treatment of skin cancer.

REFERENCES:

1. **A. M. Avery *et al.*, *Nucleic Acids Res* 27, 2256 (Jun 1, 1999).**
2. **Z. Liu *et al.*, *Proc Natl Acad Sci U S A* 108, 14831 (Sep 6).**
3. **J. Li *et al.*, *Nature* 466, 887 (Aug 12).**
4. **G. B. Sancar, A. Sancar, *Methods Enzymol* 408, 121 (2006).**
5. **B. Djordjevic, L. J. Tolmach, *Radiat Res* 32, 327 (Oct, 1967).**
6. **G. Simchen, R. Pinon, Y. Salts, *Exp Cell Res* 75, 207 (Nov, 1972).**
7. **A. B. Nuria Medina-Medina, Sean Lacey, Honghuang Lin, Elizabeth S. Klings, Clinton T Baldwin, Martin H. Steinberg, Paola Sebastiani.**
8. **K. Brettel, M. Byrdin, *Curr Opin Struct Biol* 20, 693 (Dec).**
9. **D.S. Bryan et al., Genome Res 24, 1534 (2014).**
10. **QPhotonics, LLC, 2000. Web. 6 Aug. 2014.**
11. **T. Kivioja et al., Nat Methods 9, 72 (2011).**
12. **Polidoros et al., BioTechniques 41, 35 (2006)**
13. **Hoffman et al., Nature Methods 9, 473(2012)**

APPENDICES: None

BIBLIOGRAPHY:

D. Bryan*, M. Ransom*, B. Adane, K. York, J. Hesselberth, High resolution mapping of modified DNA nucleobases using excision repair enzymes, Genome Res. 24:1534.

Ransom, M and Hesselberth, J. Poster session presented at: Cell Biology and Molecular Oncology Retreat; 2013. Oct. 21; Aurora CO.

Ransom, M and Hesselberth, J. Poster session presented at: 5th annual postdoctoral research day; 2014 Mar. 14; Aurora, CO.

Personnel receiving pay from the research:

Salary: Monica Ransom

Non-salary: None

Description of Professional Advance: During the course of this work I obtained a broad wealth of knowledge about DNA damage and repair pathways as well as sequencing library preparation and data analysis. These tools will broaden the scope of research opportunities I have as well as job availability. Big data sequencing and analysis are becoming very mainstream in research science and being able to adequately utilize and augment the available tools is becoming a must for many jobs. I have also had the opportunity to take many classes through online sources as well as at CU to further supplement my knowledge base. During my postdoctoral tenure I have had the opportunity to mentor many incoming graduate students and that mentoring has prepared me for training students in the future.